Formative Assessment: Supervised Learning

Greeshma S

•

28 Oct

10 points

Due Today

**Objective:**  
The objective of this assessment is to evaluate your understanding and ability to apply supervised learning techniques to a real-world dataset.  
  
**Dataset:**  
Use the breast cancer dataset available in the sklearn library.  
  
**Key components to be fulfilled :**  
  
**1. Loading and Preprocessing (2 marks)**

* Load the breast cancer dataset from sklearn.
* Preprocess the data to handle any missing values and perform necessary feature scaling.
* Explain the preprocessing steps you performed and justify why they are necessary for this dataset.

**2. Classification Algorithm Implementation (5 marks)**

* Implement the following five classification algorithms:  
  **1. Logistic Regression  
  2. Decision Tree Classifier  
  3. Random Forest Classifier  
  4. Support Vector Machine (SVM)  
  5. k-Nearest Neighbors (k-NN**)
* For each algorithm, provide a brief description of how it works and why it might be suitable for this dataset.

**3. Model Comparison (2 marks)**

* Compare the performance of the five classification algorithms.
* Which algorithm performed the best and which one performed the worst?

**4. Timely Submission (1 mark)**  
**Submission Guidelines:**

* Provide your code in a Jupyter Notebook format and submit the GitHub link here.
* Ensure your explanations and answers are clear and concise.

**Total Score: 10**